DOCUMENT RESUME

ED 395 236 CG 026 981

AUTHOR Gesek, Glenn J.
TITLE A Review and Critique of the Tell-Me-A-Story (TEMAS) Apperceptive Test.
PUB DATE Jan 96
NOTE 13p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, January 25-27, 1996).
PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adolescents; *Child Psychology; Cultural Differences; Diagnostic Tests; Elementary Secondary Education; Evaluation; *Evaluation Research; *Personality Assessment; *Personality Measures; *Psychological Evaluation; Psychological Testing
IDENTIFIERS *TEMAS Thematic Apperception Test

ABSTRACT

The Tell-Me-A-Story Apperceptive Test (TEMAS) is an individual apperceptive test, similar to the Thematic Apperception Test (TAT) of personality assessment. The TEMAS is intended for use with children and adolescents ages 5-18 years old, and is marketed as a personality assessment tool for minority, primarily Hispanic and African-American, and non-minority girls and boys. This paper contains a "Practical Evaluation" section, which: (1) describes the physical composition of a TEMAS; (2) discusses possible confounding factors intrinsic to the administration and interpretation of the TEMAS; and (3) describes how TEMAS stories are scored. The paper's "Technical Evaluation" section examines a sample of 642 children and discusses TEMAS' internal consistency, test-retest reliability, interrater reliability, content validity, construct validity, criterion validity, and concurrent validity as obtained from the data collected. The conclusion is drawn that the TEMAS is an excellent attempt at balancing out the disparity in multi-cultural assessment, while its primary value currently seems to be as a clinical instrument. Further studies of the African-American population's experiences with this instrument need to be conducted, and work needs to be done in general to ascertain whether the TEMAS will be suitable for suburban minority youth. (TS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 395 236

A Review and Critique of the Tell-Me-A-Story (TEMAS) Apperceptive Test

Glenn J. Gesek
Research Assistant
Department of Educational Psychology
Texas A&M University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

G. GESEK

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at the annual meeting of the
Southwest Educational Research Association,
New Orleans, Louisiana, January 25-27, 1996

BEST COPY AVAILABLE

CG026981

Test Title: TEMAS Tell-Me-A-Story (Short and Long Form)
Authors: Giuseppe Costantino, Robert G. Malgady, and Lloyd H. Rogler
Publisher: Western Psychological Services
Publication Date: 1988 (second printing in 1993)
Time Required for Administration: Short Form (9 cards): 45-60 minutes
Long Form (23 cards): 2 hours
Cost: Total- $260.00 (23 stimulus cards [minority & non-minority sets, with parallel versions for boys and girls], 25 record booklets, 1 administration instruction card, 1 technical manual).

Brief Description of Purpose and Nature of Test

The TEMAS is an individual apperceptive test, similar to the Thematic Apperception Test (TAT) of personality assessment. Intended for use with children and adolescents ages 5-18 years old, the TEMAS can be interpreted normatively for children ages 5-13 and used clinically across the entire age range. The TEMAS is marketed as a personality assessment tool for minority, primarily Hispanic and African-American, and non-minority girls and boys (Costantino, Malgady, & Rogler, 1988). The authors suggest that children tested on the TEMAS have: the cognitive ability to understand and follow instructions, the ability to devise and relate sequences of events, and willingness to cooperate. Furthermore, the TEMAS is not suggested for use with children possessing "certain sensory or neurological impairments" (Costantino et al., 1988), although this statement is not clarified further.

Unlike the TAT stimuli, the TEMAS uses colorful, unambiguous stimuli depicting situations that might be more relevant to the urban child's experience. The pictures consist of antithetical scenes, described as "bipolar personality functions" (Costantino et al., 1988) which force children to make certain specific and unconscious choices in their stories. The tester has the option of using the short or long form (9 or 23 cards). The 23 stimulus cards (11 cards have parallel versions for boys and girls) present urban settings and

themes and are based on three major premises: children identify with characters, situations, and settings that are familiar; appealing pictures capture children's interest; and realism is enhanced through color (Costantino et al., 1988).

The TEMAS is scored on personality, cognitive, and affective levels. Nine functions are scored within the personality domain: Interpersonal Relations, Aggression, Anxiety/Depression, Achievement Motivation, Delay of Gratification, Self-Concept, Sexual Identity, Moral Judgment, Reality Testing, and Functions Not Pulled (presumably omitted as a defensive reaction). As mentioned earlier, the individual cards depict antithetical scenes that force the child to make a choice in the direction that the story takes. The choice (or choices) that the child makes determines whether or not that particular personality function was "pulled" or addressed within the story. According to the authors, a few instances where the personality function is not pulled may not be a concern for the tester. Frequent omissions may indicate a defensive reaction by the child, especially if the child is repeatedly omitting the same personality function from the stories. The cognitive domain examines reaction time, total story time, fluency, omissions and transformations (character, setting, and event), sequencing, imagination, relationships, conflict, and inquiries. Stories are scored on the affective scales for happy, sad, angry, fearful, neutral, ambivalent, and inappropriate affect.

Practical Evaluation

The TEMAS is well-packaged. The technical manual is printed as a booklet, the stimulus cards are kept within sturdy protective envelopes, and the instructions fit on one laminated sheet of paper. The 142 page technical manual provides valuable information, including normative tables that are

easily understandable and accessible. While a few of the objects in the cards may be outdated (e.g.- clothing fashions, hi-fi stereos, and record players), the majority of objects, settings, and characters seem contemporary.

The TEMAS taps into a familiar child activity: story-telling. Furthermore, familiarity with the urban situations in the cards should yield richer and more meaningful stories from urban children. On the other hand, these situations might not be adequate for use with suburban minority or non-minority children. Because accurate, meaningful, and complete storytelling is threatened greatly by inadequate rapport, Costantino et al. (1988) suggest that the tester have the same ethnic background as the child, although knowledge of relevant cultural issues may be adequate.

The technical manual for the TEMAS offers no specific guides for who can and cannot use this test as an examiner. Within a sub-heading labeled "User Qualifications", the authors suggest a general knowledge of measurement, test construction methodology, and interpretive limitations. They do not, however, offer specific qualifications that one should have before using the test.

Due to the need to record stories verbatim and insert appropriate inquiries during the course of the story, computer administration and scoring of the TEMAS is not yet possible. Should children forget to address certain parts of the story, it is up to the tester to inquire about these missing parts and determine whether the child simply forgot to include them or whether he or she left them out intentionally. Furthermore, the stories might be greatly influenced by inflections, emphasis, grammatical structure, and changes made mid-sentence. The accurate measurement of such factors are not yet possible with computers. The authors also recommend recording unusual behaviors and making inquiries when parts of the story are left out or

unclear. For this reason, it may be useful to audio-tape the stories so that more attention can be paid to the child's behavior and how he or she is telling the stories.

Scoring of the TEMAS stories is detailed within the technical manual. Cognitive functions are scored dichotomously (except for Reaction Time, Total Time, and Fluency) as are Affective functions. Personality functions are rated on a 1-4 scale (1= highly maladaptive response; 4= highly adaptive response), plus N (for functions Not Pulled, explained earlier). Most cards pull for two personality functions (e.g.- the picture might deal with a situation involving Interpersonal Relations and Aggression). The manual offers some suggested responses to guide scoring, but the list is by no means definitive since individual stories can vary greatly.

Technical Evaluation

The TEMAS standardization sample consisted of 642 children from public schools in the New York City area: 281 males, 361 females. Of those children, 172 were White, 206 African-American, 164 Puerto Rican, and 100 "other Hispanic". Though large in number, this sample is generalizable only to New York City minority and non-minority children, not necessarily to children across the nation. Although the normative data presented in the manual are based on an unstratified sample and should be considered preliminary data (Costantino et al., 1988), a better sample would have considered geographical distribution, socioeconomic level, and relevant personality characteristics (Anastasi, 1988).

Although the sample was not stratified at the outset, normative data are stratified by ethnic group and age: 5-7, 8-10, and 11-13 years of age. A possible problem exists within the youngest age range: 5-7 years old. Developmentally, the 5-7 shift is often one of substantial cognitive gain. At

this stage of development, the child is capable of more complex thought and understanding (such as Piaget's concepts of Conservation and Perspective-Taking). Placing the scores of 5- and 7-year-olds within one group may be mixing children with two vastly different levels of cognitive ability, thereby granting a different picture of those children's abilities.

Ethnically, recent work (not yet included in the technical manual) extended the norms for the TEMAS across three Hispanic Subcultures: native Puerto Ricans, New York Puerto Ricans, and Argentineans (Costantino, Malgady, Casullo, & Castillo, 1991). Such normative extension has not yet been examined with African-American or non-minority children. Is one sample of African American children from New York City adequate to determine that the use of the TEMAS is both reliable and valid for this ethnic group across the country? When one considers the differences between geographical regions in this country, the answer to the above question seems quite obvious. Furthermore, by limiting themselves to the population used, the authors are defining "minority" as urban children, excluding suburban or rural children across the country (even the more recent study only tests urban children). The normative base, and its limits, suggests the usefulness of the TEMAS as a clinical instrument for all ethnic groups, but not as a normative instrument measuring personality for African-Americans or suburban minorities.

When scoring the TEMAS for normative comparison or clinical use, quantitative scores are transformed into T scores (Mean 50, Standard Deviation 10), while qualitative scores are based on critical cutoffs (using the 90th percentile). Qualitative, in this sense, does not refer to how the data are analyzed, but rather the type of data that are being analyzed. Qualitative scores exist only within the affective and cognitive functions, not the

personality domain. Furthermore, whereas Reaction Time and Story Length are quantitative scores in the cognitive domain, Imagination and Relationships (identified or not by the child) are considered qualitative; primarily because the tester has more leeway in the interpretation of their presence or absence.

Short form data were obtained by extracting the scores from those cards given in the long form. Clinicians should be careful when interpreting short form scores: factors such as test fatigue may have affected children's performances on the long form that would not be a factor when given the short form alone. Correlations are reported between short and long forms within the technical manual (.81 for Total sample, .82 for Whites, .80 for African-Americans and Hispanics). However, such high correlations are heavily influenced by the fact that the cards in the short form are included in the total score of the long form, thus including 9 correlations of 1.00 into the calculations.

The TEMAS technical manual also offers internal consistency, test-retest, and interrater reliabilities for both short and long forms. Internal consistency coefficients for the Hispanic sample (N=73) ranged from .41 (Ambivalent Affect) to .98 (Fluency), with the median of .73. Coefficients for the African-American sample (N=42) ranged from .31 (setting transformations) to .97 (fluency)- median .62- with 13 out of 32 measures at or above .70. In the personality domain, alphas for the African-American sample were uniformly lower than those of the Hispanic sample; the highest reliabilities were obtained for Aggression and Moral Judgment. With a difference of 30 children between the samples, the "uniformly lower alphas" may not be as problematic as they appear, or they might be heightened with more equivalent sample numbers. Using a separate sample of 210 Puerto Rican

children (Mean Age= 7.45 years, SD 1.53), Short Form alpha coefficients ranged from .30 (fearful affect) to .92 (fluency and total time), median .74.

Test-retest reliability was measured taking 51 of the 210 Puerto Rican children from the Short Form internal consistency sample and testing them again. It is surprising that there was not a similar sample of African American children drawn and retested as well. After 18 weeks, the most significant test-retest value (according to the authors) was .53 for Conflict Scores in the Cognitive domain. Other statistically significant values reported ranged from .35 for Happy Affect to .46 for Event Transformations. Oddly, this was the only table in the chapter to report p values (hence the report of "statistically significant values"), causing one to question whether the authors were attempting to make problematic data look better.

Aside from the questionable report of p values, these values are fairly low (.35 being significant explains 12.25% of the variance), and may be as much an artifact of the sample size as anything else. Possible explanations (by the authors) for such low test-retest scores include: inconsistent subjects, the use of different raters, limited range of indicators from the instrument, and the possibility of less stable scores from children who exhibit behavior problems (Costantino et al., 1988). Two other alternatives do exist. Perhaps the children just noticed different parts of the picture the second time around (thus leading the authors to call *them* "inconsistent" when the pictures may be too complex), or perhaps the TEMAS is not yet as strong a test as the authors claim.

Interrater reliability was determined by independent ratings of randomly drawn children, 27 Hispanic and 26 African-American, from the Long Form internal consistency sample. Interrater reliabilities were consistently moderate to high for Hispanics and African-Americans (ranging from .27-

1.00). No medians were reported, although tables for both samples of children were included in the technical manual for the scores in each of the three domains.

Content, construct, and criterion validity were also addressed in the technical manual. Agreement between practicing school and clinical psychologists ranged from 71-100%, across pictures, concerning content validity; i.e. whether the pictures were really focusing on the personality functions intended by the authors. Within this study, the psychologists were given the list of 9 personality functions on the test and asked to determine which functions were pulled by the pictures.

Construct validity was evaluated by comparing 210 Puerto Rican students rated below the median on a variety of classroom behaviors (from aggression to anxiety) to the general Hispanic sample scores used for internal consistency ratings. Behaviorally screened children responded to inquiries less often; reduced fluency while increasing total time; exhibited more frequent omissions and transformations; tended less often to attribute happiness, anger, and fear to the characters; and consistently rated lower for adaptiveness on the personality functions than the general population (Costantino et al, 1988).

Concurrent validity was assessed using the same sample. To determine how TEMAS results related to ego development, they were compared with results of the Sentence Completion Test. A similar comparison was made with teacher behavior ratings as determined by the Teacher Rating Behavior Scale. Graduate Psychology students served as observers and independently rated the children regarding: delay of gratification, self-concept, disruptive behavior, and aggression. Regression analyses are reported within the test manual, with significant F values found between the TEMAS and: ego

development ($F(6, 116)=3.47$, $p<.05$), teachers' behavior ratings ($F(12,102)=2.69$, $p<.05$), delay of gratification ($F96,116)=2.21$, $p<.05$), self-concept ($F(15,107)=2.38$, $p<.05$), disruptive behavior ($F(15,107)=2.51$, $p<.05$), and aggression ($F(6,116)=2.21$, $p<.05$). The only non-significant value was found between the TEMAS and trait anxiety (using the State-Trait Anxiety Inventory for Children).

Reviewer Comments

Reviews published previously by Cambias, Killian, and Faust (1992) and Ritzler (1993) give favorable evaluations of the utility of the test, while questioning the usefulness for African-American populations. Ritzler warns against interpretation with rural or higher-social-class children, further emphasizing caution for African-American generalization; Lang (1992) also agrees on this point. In all three reviews, questions are raised regarding reliability values, especially test-retest, although Cambias et al. (1992) praise the TEMAS as "a serious attempt to address the lack of psychometric rigor characteristic of other apperceptive tests" (pg. 557).

Lang (1992), however, does not consider the TEMAS to be an improvement in apperceptive testing, and believes that the TAT is still better for testing minority children. Though psychometrically justifiable at this point, Lang's statement could simply be due to the longer history of use with the TAT and the more rigorous development and study of that test. Practically, pictures that offer a sense of familiarity for children seem preferable to pictures that depict non-minority characters in unfamiliar settings. Lang's point is well-taken, in that failure to investigate and report concurrent validity with the TAT within the technical manual may be a shortcoming of the TEMAS development.

Summary Evaluation

The TEMAS is an excellent attempt at balancing out the disparity in multi-cultural assessment. At this point, its primary value seems to be as a clinical instrument, though generalizability using norms is possible with Hispanic populations. Low reliabilities are a primary concern (especially test-retest) and may require a scoring adjustment or omission of certain inconsistently addressed personality functions (suggested also by Ritzler, 1991). The use of colorful unambiguous stimuli seems to be a step in the right direction, but in order for this to be a tool for *minority* assessment, further work needs to be done with the African-American population. For instance, many aspects of reliability for the African-American samples were not tested by the authors. A study of test-retest reliability needs to be done, as does a study extending the sample to other urban and even suburban African-American children. Concurrent and predictive validity also needs to be assessed, in the same manner as has been done but utilizing a sample of African-American children. Furthermore, work needs to be done in general to ascertain whether the TEMAS will be suitable for suburban minority youth, for whom current tests may be equally invalid, but also for whom the urban experiences depicted by the TEMAS might not be adequate. It is hoped that such work is on the agenda for the test authors, or else they will need to clarify the advertisements that promote the TEMAS as a multi-cultural apperception test and specify it instead for urban Hispanic populations.

References

Anastasi, A. (1988). Psychological Testing (6th ed.). New York: Macmillan.

Cambias, R. D. (Jr.), Killian, G. A., & Faust, J. (1992). Review of TEMAS (Tell-me-a-story). In Test Critiques Vol. IX. Austin, TX: Pro-Ed.

Costantino, G., Malgady, R. G., Casullo, M. M., & Castillo, A. (1991). Cross-cultural standardization of TEMAS in three Hispanic subcultures. Hispanic Journal of Behavioral Sciences, 13(1), 48-62.

Costantino, G., Malgady, R. G., & Rogler, L. H. (1988). TEMAS (Tell-me-a-story) Manual. Los Angeles: Western Psychological Services.

Lang, W. S. (1992). Review of the TEMAS. In J.J. Kramer & J. C. Conoley (Eds.) The Eleventh Mental Measurements Yearbook. Lincoln, Nebraska: University of Nebraska Press.

Ritzler, B. (1993). Test Reviews: the TEMAS. Journal of Psychoeducational Assessment, 11, 381-389.